Summary Report of North American Industrial Classification System and Standard Industrial Classification Data Standards

CONTRACT NO. 68-WI-0055

DELIVERY ORDER NO. 57

Prepared for:

United States Environmental Protection Agency

Office of Information Resources Management

401 M Street SW

Washington, DC 20460

Delivery Order Project Officer:

Geoffrey Steele

Prepared by:

EPA Systems Development Center

(A Contractor Operated Facility)

Science Applications International Corporation

200 North Glebe, Suite 300

Arlington, VA 22203

CONTENTS

SDC-0055-057-LF-6028
August 29, 1997

1.0 INTRODUCTION

The Environmental Protection Agency (EPA) Office of Information Resources Management (OIRM) Enterprise Information Management Division (EIMD), through the mission of its Information and Data Management (IDM) Program, has introduced the concept of centralized management and coordination of EPA's distributed information and data resources. The data policies and standards of the IDM program will be supported by a data architecture, data models, standard data element domains, and a repository of standard data elements. These work products will facilitate data sharing efforts and demonstrate EPA's effective management of its Information Resources Management (IRM) resources.

1.1 Background

The One Stop Reporting Program was initiated by the EPA Executive Steering Committee (ESC) for the purposes of improving the Agency's management of environmental information and reducing the burden of reporting across the regulated community. The intended process for improvement of information management is to involve the stakeholders, particularly the states, in identifying data elements and domains that need to be standardized. To further the goals of the One Stop Program, the Information and Data Management Service Center (IDMSC), Delivery Order (DO) 57 has been tasked to identify the need for data standards that will facilitate data sharing and to develop data standards that are used within the Agency and its state partners. One need identified is the development of standard data elements and domains for representation of the economic activities of business establishments.

A common method of classifying businesses or industries by type is the 1987 Standard Industrial Classification (SIC) System, commonly referred to as the SIC Code. It was developed by the United States (U.S.) Government in conjunction with U.S. businesses. Many databases use the SIC number as an index. It divides virtually all economic activity into divisions that are further broken up into numbered major groups. The reference book is the Standard Industrial Classification Manual 1987, prepared by the Executive Office of the President, Office of Management and Budget and available through the National Technical Information Service (NTIS) and Government bookstores. It is in the process of revision to better account for the changing structure of modern business.

The Office of Management and Budget (OMB) established an Economic Classification

Policy Committee (ECPC) which was charged with developing a North American Industry Classification System (NAICS) for 1997. OMB charged the ECPC with a "fresh slate" examination of economic classifications for statistical purposes, including industrial classifications, product classifications, and product code groupings. The classification was developed in cooperation with Statistics Canada and Mexico's Instituto Nacional de Estadìstica, Geografia e Informatìca (INEGI). Statistics Canada and INEGI have accepted the new NAICS system, which will be used for industrial classification in the statistical programs of Canada and Mexico. The Executive Office of the President, published the final recommendations for *Economic Classification Policy Committee: Standards Industrial Classification Replacement--The North American Industry Classification System Proposed Industry Classification Structure* on November 5, 1996, with the intention of replacing the SIC codes with NAICS codes on January 1, 1997.

1.2 Purpose

The purpose of this report is to document the standard data elements and domains developed for SIC and NAICS and to certify the authority for their accuracy and validity.

1.3 Scope

This document is limited to the standard data elements that have been registered in the EPA Environmental Data Registry (EDR) and the corresponding domains of valid data for those data elements.

1.4 References

The following reference materials were used in the preparation of the SIC and NAICS standard data elements and domain values:

• *Standard Industrial Classification Manual 1987*, Executive Office of the President, Office of Management and Budget.

• *Economic Classification Policy Committee: Standards Industrial Classification Replacement--The North American Industry Classification System Proposed Industry Classification Structure*, Executive Office of the President, November 5, 1996.

• *1987 SIC Manual*, electronic copy, U.S. Commerce Department.

• NAICS-SIC Crosswalk Table 1, U.S. Commerce Department, (http://www.census.gov/epcd/www.naicstab.htm).

2.0 METHODOLOGY

The Project Team followed the One Stop data standards development procedure. Data standards include two components: standard data elements recorded in the EDR and standard data values (i.e., domains) for all data elements where only a discrete set of data values are acceptable. The procedure followed included analysis of the requirements, development of a logical data model in third normal form, recording the standard data elements in the EDR, identification of an authoritative source of domain data, preparing that data in the appropriate formats for update to the EDR, and update of that domain data into the EDR database.

2.1 Data Model for Industrial Classifications

The data model for industrial classifications, both SIC and NAICS, is provided in Appendix A.

2.2 Description of Standard Industrial Classifications

The SIC is divided into divisions (encoded A through K) which are subdivided into major groups (2-digit codes represented as MM). The SIC codes are hierarchical. The 3-digit industry group codes (represented as MMG) are grouped within the major groups, as are the 4-digit industrial classification codes (represented as MMGI). Additionally, an uncoded set of industry subdivisions provides further breakdown of the 4-digit codes. Organizations can establish codes for those subdivisions, but there are no standard codes. Most business establishments, however, use the 4-digit codes. Industries are often indexed by the 2-digit major groups for the purposes of generating statistics and to associate related industries.

2.3 Standard Data Elements for Standard Industrial Classifications

The standard data elements (DE) developed for SIC are the following:

Industrial Classification Division Code -- The code that represents a division in the economy that covers an economic activity as specified by the SIC. (DE 5861, 1-char).

Industrial Classification Major Group Code -- The code that represents a major group of industrial classifications as specified by the SIC. (DE 5869, 2-char [MM]).

Industrial Classification Group Code -- The code that represents a group of related industries within the economy as specified by the SIC. (DE 5865, 3-char [MMG]).

Standard Industrial Classification Code -- The code that represents the economic activity of a company as specified by the SIC. (DE 5338, 4-char [MMGI]).

Industrial Classification Subdivision Text -- The text that describes a specific component of an industry as specified by the SIC. (DE 5867, 96-char).

Detailed information about these data elements, their domains, and permissible values are provided in Appendix B as reports generated from the EDR.

2.4 Description of the North American Industrial Classification System

The NAICS is organized similarly to the SIC codes. Economic activities of business establishments are grouped hierarchically into economic sectors (E), economic subsectors (S), industry groups (G), and industries (I). Industry classifications are further subdivided into national classifications that are specific to the needs of the U.S. (N). The economic sectors are represented as 2-digit codes (EE); the economic subsectors as 3-digit codes (EES); the industry groups as 4-digit codes (EESG); the industries as 5-digit codes (EESGI); and the U.S. industry national subdivisions as 6-digit codes (EESGIN).

2.5 Standard Data Elements for North American Industrial Classification System

The standard data elements (DE) developed for NAICS are the following:

• North American Industrial Economic Sector Code -- The code that represents an economic sector as specified in the NAICS. (DE 5847, 2-char [EE]).

• North American Industrial Classification Subsector Code -- The code that represents an economic subsector as specified in the NAICS. (DE 5853, 3-char [EES]).

• North American Industrial Classification Group Code -- The code that represents a group of related industries within the economy as specified in the NAICS. (DE 5849, 4-char [EESG]).

• North American Industrial Classification Code -- The code that represents an industry within the economy as specified in the NAICS. (DE 5851, 5-char [EESGI]).

• North American U.S. National Industrial Code -- The code that represents a subdivision of an industry that accommodates user needs in the U.S. as specified in the NAICS. (DE 5855, 6-char [EESGIN]).

Detailed information about these data elements, their domains, and permissible values are provided in Appendix C as reports generated from the EDR.

3.0 WORLD WIDE WEB LINKAGES

As U.S. information systems convert from the SIC to the NAICS codes for classification of business economic activities, it is necessary that a cross reference of SIC to NAICS codes be available to users. A one-to-one relationship between SIC and NAICS codes does not exist. For example, a 4-digit SIC code might describe the same industrial classification as a 5- or 6-digit NAICS. There are several sites on the World Wide Web that provide search functionality. There are also cross reference tools for using the SIC and NAICS information that will be helpful for converting SIC to NAICS. The WWW also provides access to more information about these systems of industrial classification. These sites include:

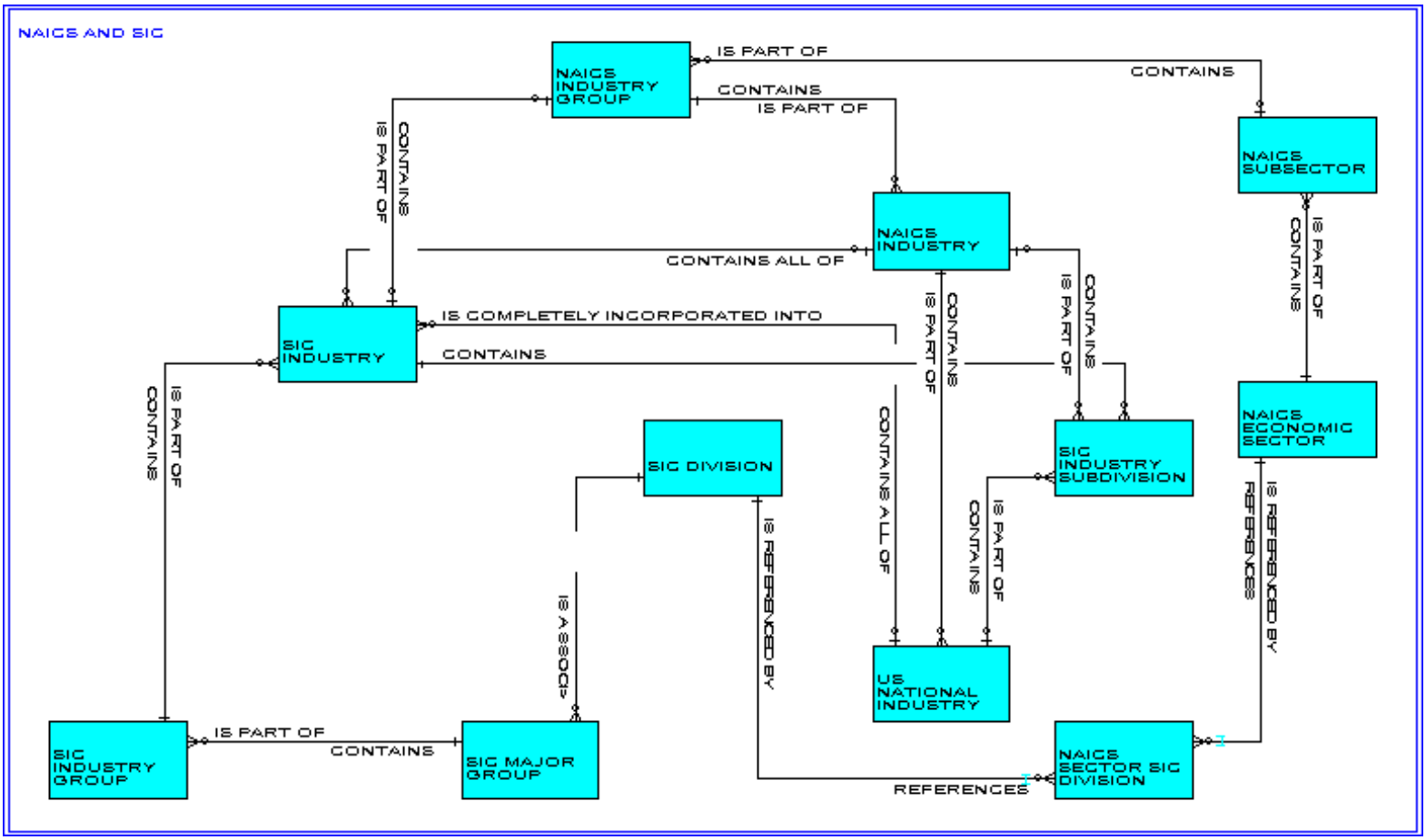
• http://www.osha.gov/oshstats/sicser.html -- The U.S. Department of Labor, Office of Safety and Health Administration (OSHA) site provides a good search tool for searching SIC by code or text. The OSHA site has linkages to NTIS and to an overview of the SIC.

• http://www.census.gov/epcd/www/naicstab.htm -- The U.S. Department of Commerce, Bureau of the Census site provides cross references for SIC to NAICS information as well as the reverse.

• http://www.vivamus.com/sic.html -- Vivamus Concepts provides a site with a complete listing of SIC codes with their hierarchies.

• http://http.tamu.edu/~jennyb/naics.html -- This site is an unofficial clearinghouse about NAICS. It contains references to publications and other information about NAICS and its usage.

APPENDIX A Data Model for SIC and NAICS Information



Data Model for SIC and NAICS Information

APPENDIX B Standard Data Elements for Standard Industrial Classification Data

The information contained in Appendix B can be obtained from the Environmental Protection Agency's Environmental Data Registry. To view the current information, select **Data Element Identifier and Version Number Query** and enter the Data Element Identifier and Version Number for the elements listed in the table below.

ID	Version	Name
5861	1	Industrial Classification Division Code
5869	1	Industrial Classification Major Group Code
5865	1	Industrial Classification Group Code
5338	1	Standard Industrial Classification Code
5867	1	Industrial Classification Subdivision Text

APPENDIX C Standard Data Elements for North American Industrial Classification Data

The information contained in Appendix C can be obtained from the Environmental Protection Agency's Environmental Data Registry. To view the current information, select **Data Element Identifier and Version Number Query** and enter the Data Element Identifier and Version Number for the elements listed in the table below.

ID	Version	Name
5847	1	North American Industrial Classification Economic Sector Code
5853	1	North American Industrial Classification Subsector Code
5849	1	North American Industrial Classification Group Code
5851	1	North American Industrial Classification Code
5855	1	North American Industrial Classification Subdivision Text

DRAFT